Kepler Data Release 12 Notes

KSCI-19052-001
Data Analysis Working Group (DAWG)
Thomas Barclay (Editor)
Jessie Christiansen (Editor)

Data Release 12 for Quarter Q9

Q.m		First Cadence MJD midTime	Last Cadence MJD midTime	First Cadence UT midTime	Last Cadence UT midTime	Num CINs	Start CIN	End CIN
9	LC	55641.016958	55738.423953	21-Mar-2011 00:24:25	26-Jun-2011 10:10:29	4768	34237	39004
9.1	SCM1	55641.007082	55677.419086	21-Mar-2011 00:10:11	26-Apr-2011 10:03:28	53460	1015570	1069029
9.2	SCM2	55678.114509	55706.618707	27-Apr-2011 02:44:53	25-May-2011 14:50:56	41850	1070050	1111899
9.3	SCM3	55707.252830	55738.433829	26-May-2011 06:04:04	26-Jun-2011 10:24:42	45780	1112830	1158609

Prepared by: ______________________ Date: 11-30-2011
Thomas Barclay, Kepler Science Office, for the DAWG (next page)

Approved by: ______________________ Date: 11/30/11
Jeffrey Van Cleve, Editor-in-Chief

Approved by: ______________________ Date: 11-30-11
Jon Jenkins, Co-I for Data Analysis & DAWG Lead

Approved by: ______________________ Date: 11/30/11
Michael R. Haas, Kepler Science Office Director

Document Control

Ownership
This document is part of the Kepler Project Documentation that is controlled by the Kepler Project Office, NASA/Ames Research Center, Moffett Field, California.

Control Level
This document will be controlled under KPO @ Ames Configuration Management system. Changes to this document **shall** be controlled.

Physical Location
The physical location of this document will be in the KPO @ Ames Data Center.

Distribution Requests
To be place on the distribution list for additional revisions of this document, please address your request to the Kepler Science Office:

Michael R. Haas
Kepler Science Office Director
MS 244-30
NASA Ames Research Center
Moffett Field, CA 94035-1000
Michael.R.Haas@nasa.gov

The Data Characteristics Handbook and accompanying Data Release Notes are the collective effort of the Data Analysis Working Group (DAWG), composed of Science Office (SO), Science Operations Center (SOC) and Guest Observer Office (GO) members as listed below:

Jon Jenkins, Chair
Doug Caldwell, Co-Chair
Allen, Christopher L.
Barclay, Thomas
Bryson, Stephen T.
Burke, Christopher J.
Christiansen, Jessie L.
Clarke, Bruce D.
Cote, Miles T.
Fanelli, Michael N.
Girouard, Forrest
Haas, Michael R.
Hall, Jennifer
Ibrahim, Khadeejah
Kinemuchi, Karen
Klaus, Todd
Kolodziejczak, Jeffery (MSFC)
Li, Jie
McCauliff, Sean D.
Middour, Christopher K.
Morris, Rob
Mullally, Fergal
Quintana, Elisa V.
Rowe, Jason
Seader, Shawn
Smith, Jeffrey Claiborne
Still, Martin
Stumpe, Martin C.
Tenenbaum, Peter G.
Thompson, Susan E.
Twicken, Joe
Uddin, Akm Kamal
Van Cleve, Jeffrey
Wohler, Bill

Contents

A Word of Caution to Users

The Simple Aperture Photometry (SAP) product generated by the PA (Photometric Analysis) pipeline module and the PDC (Pre-search Data Conditioning) processed data are designed for automated photometry on over 160,000 stars. Although significant effort has been expended to make these products robust and reliable, they do not produce the best photometry for every target. To mitigate the impact of non-optimal apertures, Kepler is now providing target pixel files so that users can perform their own photometry. To mitigate known issues with PDC, Kepler has completely reworked this software module and is now producing significantly improved results for most long-cadence targets (see Appendix A.1). To take advantage of these improvements, all quarters of Kepler data will be reprocessed and redelivered to MAST by late 2012. Meanwhile, cotrending basis vectors (see Section 7.0) have been provided for all past quarters so that users can better remove systematics from the simple-aperture-photometry (SAP) light curves generated by the pipeline or custom light curves self-extracted from the target pixel files.

Investigators are strongly encouraged to study the Data Characteristics Handbook and Data Release Notes for all data sets that they intend to use. We advise against publication of results based on Kepler light curves without careful consideration and due diligence by the end user, and dialog with the Science Office or Guest Observer Office where appropriate.

Users are encouraged to notice and document artifacts, either in the SAP or PDC data, and report them to the Science Office at kepler-scienceoffice@lists.nasa.gov.

Like Waldseemüller's 1507 map, the Maximum A Priori (MAP) method for PDC and other improvements to light-curve correction implemented in Kepler Pipeline 8.0 will show us more of the light-curve world than we knew before, while leaving some hazards to navigation around the edges. Far-ranging light-curve mariners will be given the tools to help navigate in such perilous circumstances by the Guest Observer Office.

1 Introduction

These Data Release Notes provide information specific to the release of Q9 data, processed with SOC Pipeline 8.0. These Notes contain the summary figures and tables for this quarter—the companion text can be found in the Kepler Data Characteristics Handbook (KSCI-19040). The sections are numbered in the same order in these Notes and the Handbook to assist the reader, except that Section 5.15 in this document is Section 5.14 in the Data Characteristics Handbook due to the addition of a new Section 5.14.

1.1 Dates, Cadence Numbers, and Units

No changes from the Data Characteristics Handbook.

Contents of Data Release 12.

Q.m		First Cadence MJD midTime	Last Cadence MJD midTime	First Cadence UT midTime	Last Cadence UT midTime	Num CINs	Start CIN	End CIN
9	LC	55641.016958	55738.423953	21-Mar-2011 00:24:25	26-Jun-2011 10:10:29	4768	34237	39004
9.1	SCM1	55641.007082	55677.419086	21-Mar-2011 00:10:11	26-Apr-2011 10:03:28	53460	1015570	1069029
9.2	SCM2	55678.114509	55706.618707	27-Apr-2011 02:44:53	25-May-2011 14:50:56	41850	1070050	1111899
9.3	SCM3	55707.252830	55738.433829	26-May-2011 06:04:04	26-Jun-2011 10:24:42	45780	1112830	1158609

2 Release Description

No changes from the Data Characteristics Handbook.

3 Evaluation of Performance

3.1 Overall

The 6.5-h Temporal Median (TM) of the Q9 CDPP time series calculated by the Transiting Planet Search (TPS) pipeline is a measure of signal to noise. Compared to Q8, we see a significant reduction in the median of the value calculated for stars brighter than Kp=12. We do not see a similar decrease when considering only the dwarf stars and we conclude that the quality of giant star light curves has improved dramatically. We attribute this to the new implementation of the Pre-search Data Conditioning pipeline module (see Appendix A.1).

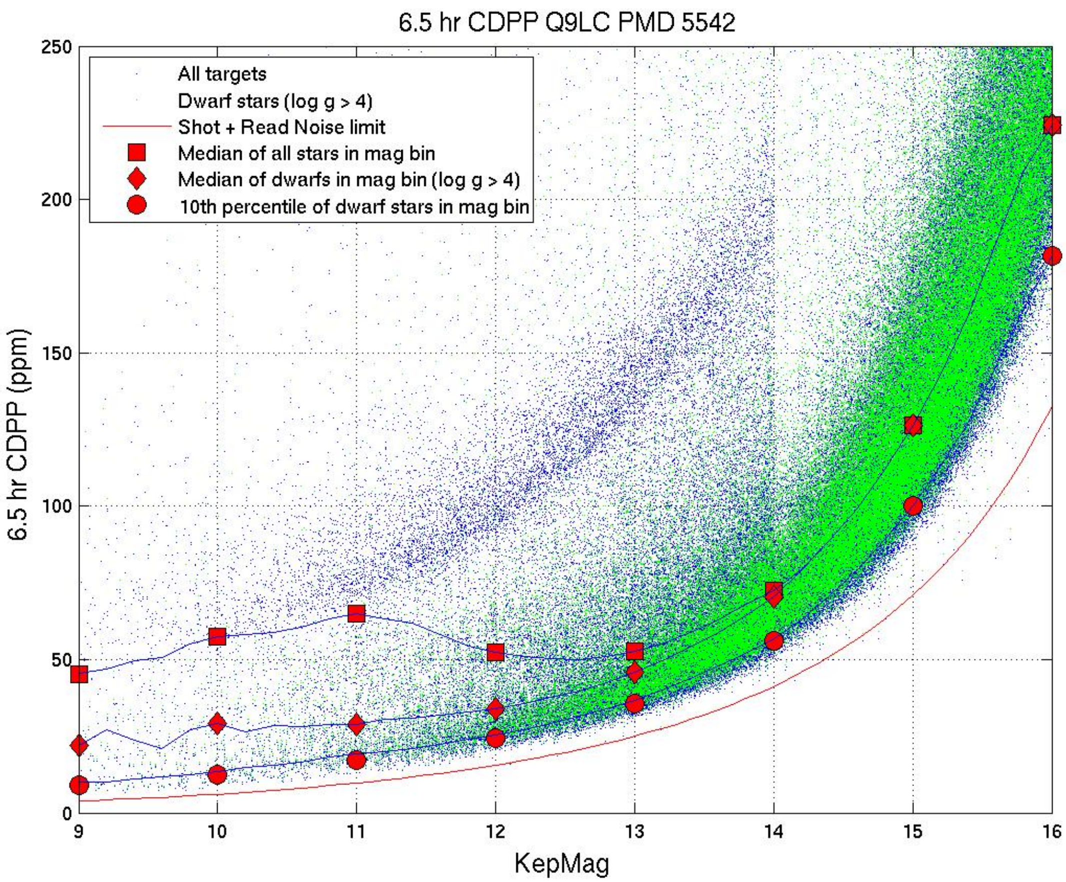

Figure 1: 6.5-h Temporal Median (TM) of the Q9 CDPP time series calculated by the TPS pipeline module for stars between Kp=9–16. The 6-h TMCDPPs have been divided by sqrt(13/12) = 1.041 to approximate 6.5-h TMCDPPs. Stars on the planetary target list with log g > 4, which are likely to be dwarf stars, are shown as green + symbols; other stars are marked with blue + symbols.

Table 1: Aggregate statistics for the TMCDPPs plotted in Figure 1. Column Definitions: (1) Kepler Magnitude at the center of the bin. Bins are ± 0.25 mag, for a bin of width 0.5 mag centered on this value. (2) Number of dwarfs (log g > 4) in the bin. (3) 10th percentile TMCDPP for dwarfs in the bin. (4) Median TMCDPP for dwarfs in the bin. (5) Number of all stars in the bin. (6) 10th percentile TMCDPP of all observed stars in the bin. (7) Median TMCDPP for all stars in the bin. (8) Simplified noise model CDPP.

Kp mag	No. dwarfs	10th prctile	Median	No. stars	10th prctile	Median	Noise model
9.0	52	9.2	21.9	185	10.2	45.4	3.8
10.0	161	12.5	29.2	591	13.3	57.3	6.0
11.0	635	17.4	28.8	1776	19.2	64.9	9.5
12.0	2229	24.3	34.0	4291	25.5	52.4	15.2
13.0	7009	35.8	46.0	10071	36.7	52.6	24.4
14.0	14315	56.2	70.3	16545	56.7	72.3	40.1
15.0	28732	100.1	126.3	28736	100.1	126.3	68.8
16.0	15020	181.6	224.2	15020	181.6	224.2	127.8

4 Historical Events

In this Section, we discuss cadences that may not be useful for high-precision photometry due to planned or unplanned spacecraft events.

4.1 Kepler Mission Timeline to Date

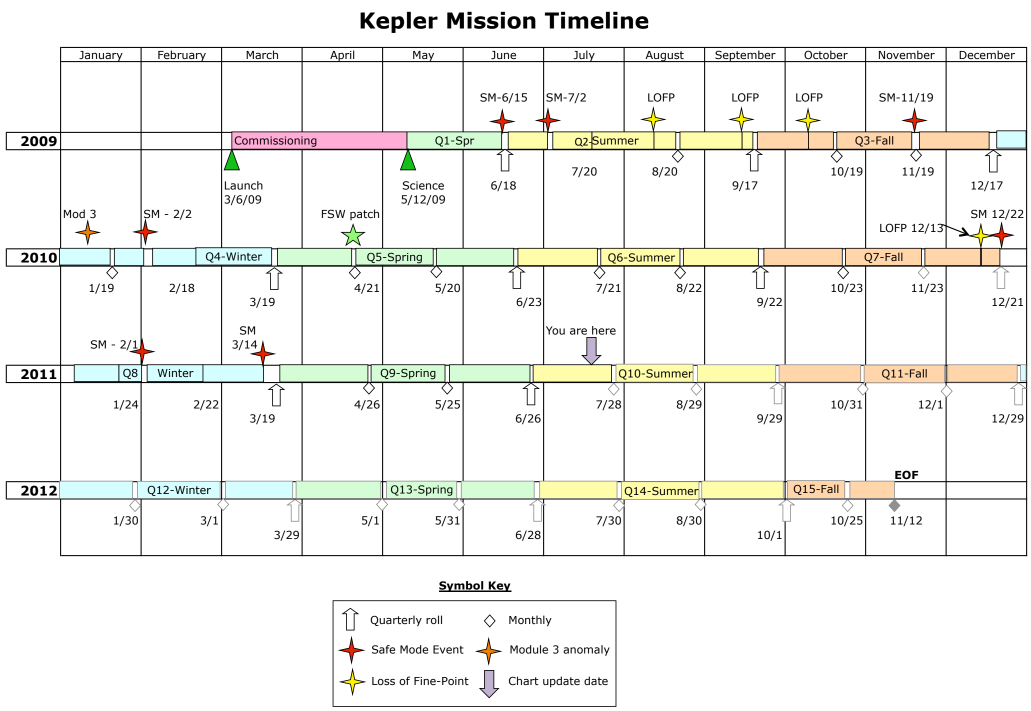

Figure 2: Kepler Mission Timeline as of July 2011.

4.2 Safe Mode

There were no Safe Modes in Q9. The Safe Mode shown in Figure 2 on 2011-03-14 occurred before the beginning of the quarter and has no impact on the Q9 data.

4.3 Loss of Fine Point

The cadences obtained when the spacecraft was not in fine point are listed in Tables 8 and 9, the LC and SC anomaly summary tables respectively.

4.4 Attitude Tweaks

The pointing history is shown in Figure 3; there were no attitude tweaks during Q9.

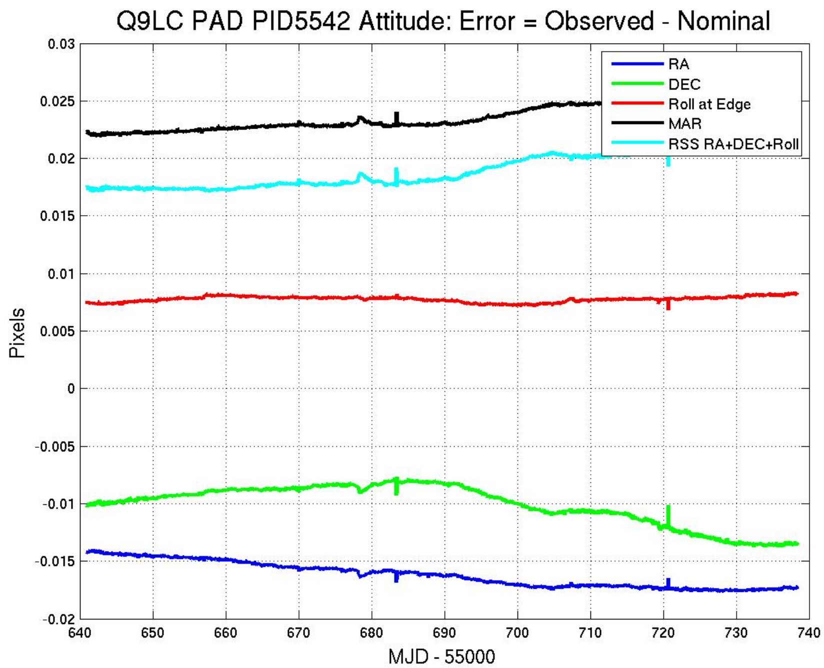

Figure 3: Attitude Error in Q9, calculated by PAD (Photometer Attitude Determination) using Long Cadence data.

4.5 Variable FGS Guide Stars

No changes from the Data Characteristics Handbook.

4.6 Module 3 Failure

No changes from the Data Characteristics Handbook.

5 Ongoing Phenomena

In this Section, we document the systematic errors arising in nominal on-orbit operations, most of which will be removed from the PDC flux time series by the science pipeline.

5.1 Image Motion

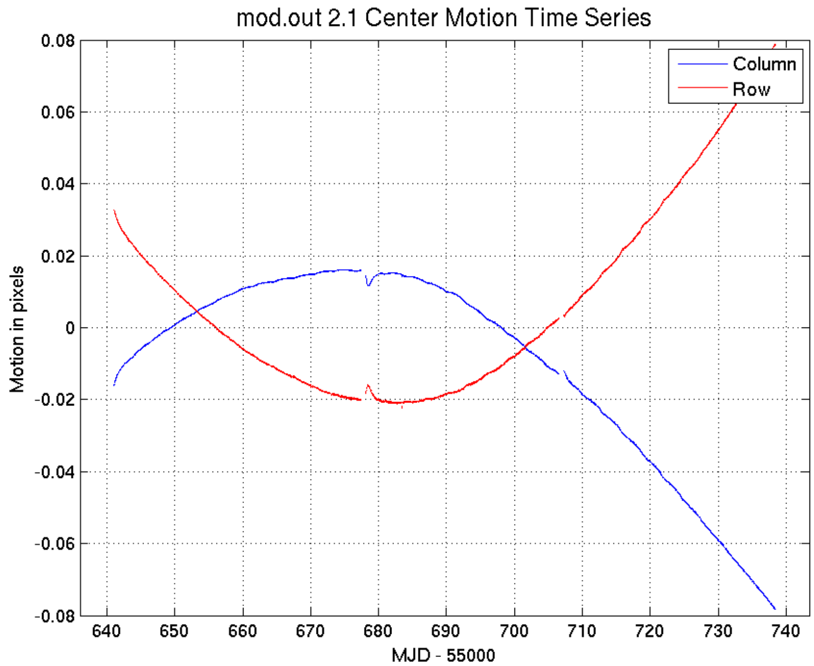

Figure 4: The center motion time series for module 2, output 1. The gaps at MJD-55000 = 678 and 707 are monthly Earth contacts.

The image motion on a per-target basis is now available in the FITS files at MAST via the POS CORR1 (column) and POS CORR2 (row) columns. In Figure 4 we show the column and row motion time series for the center of mod.out 2.1; these time series are provided for each mod.out in the Supplement.

5.2 Focus Changes

The change in width of the PRF with time is shown in Figure 5.

5.3 Momentum Desaturation

The Long Cadences affected by momentum desaturations in the reaction wheels are listed in Table 2, and the Short Cadences in Table 3. They are also flagged in the FITS files at MAST in the quality flag column (see the Kepler Archive Manual).

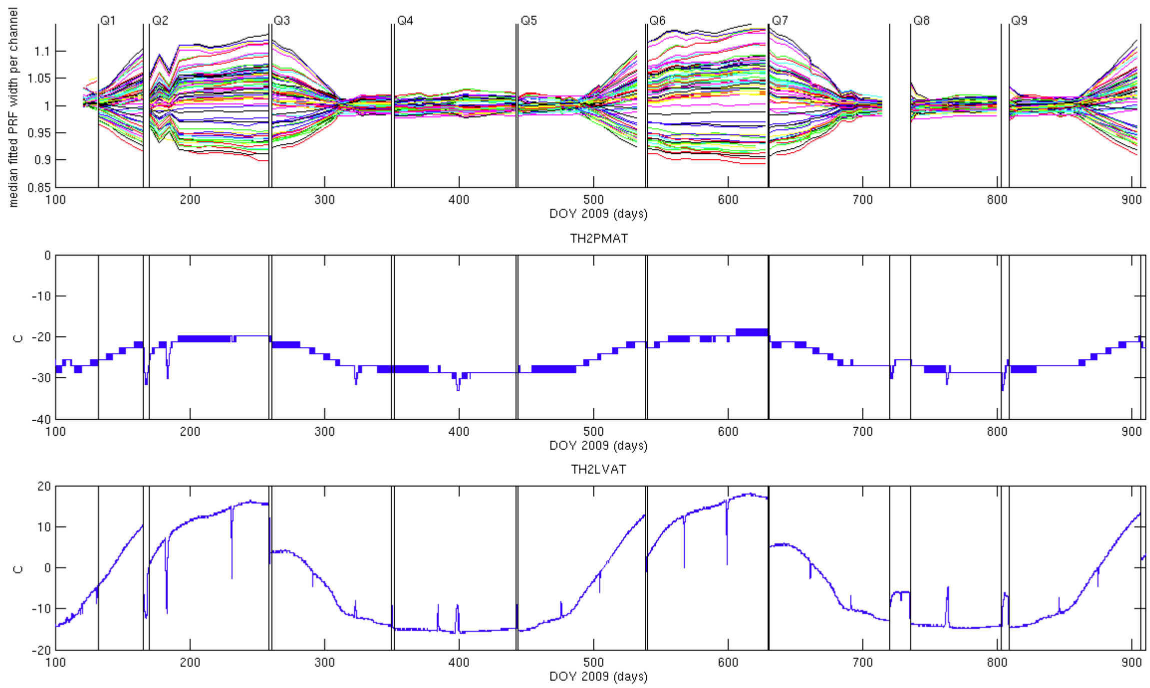

Figure 5: The top plot shows the variation of PRF width with time. The middle and lower plots show the temperature variations of the primary mirror and the launch vehicle assembly. The correlation of the spacecraft temperature and PRF widths demonstrates the seasonal nature of the focus.

Table 2: Momentum desaturations in Q9 and the corresponding Long Cadences. CIN = cadence interval number, RCI = relative cadence index.

LC		
CIN	RCI	Date (MJD)
34340	104	55643.12162
34487	251	55646.12536
34633	397	55649.10867
34779	543	55652.09197
34925	689	55655.07528
35071	835	55658.05859
35217	981	55661.04189
35363	1127	55664.02520
35509	1273	55667.00850
35655	1419	55669.99181
35801	1565	55672.97512
35947	1711	55675.95842
36166	1930	55680.43338
36312	2076	55683.41669
36458	2222	55686.40000
36604	2368	55689.38330
36750	2514	55692.36661
36896	2660	55695.34991
37042	2806	55698.33322

Continued on next page

Table 2 – continued from previous page

CIN	RCI	Date (MJD)
37188	2952	55701.31653
37334	3098	55704.29983
37588	3352	55709.48997
37734	3498	55712.47328
37880	3644	55715.45658
38026	3790	55718.43989
38172	3936	55721.42319
38318	4082	55724.40650
38464	4228	55727.38981
38610	4374	55730.37311
38756	4520	55733.35642
38902	4666	55736.33973

Table 3: Momentum desaturations in Q9 and the corresponding Short Cadences. CIN = cadence interval number, RCI = relative cadence index. The months are separated by horizontal lines.

SC		
CIN	RCI	Date (MJD)
1018670	3101	55643.11855
1018671	3102	55643.11924
1023080	7511	55646.12229
1023081	7512	55646.12298
1027460	11891	55649.10560
1027461	11892	55649.10628
1031840	16271	55652.08891
1031841	16272	55652.08959
1036220	20651	55655.07221
1036221	20652	55655.07289
1040600	25031	55658.05552
1040601	25032	55658.05620
1044980	29411	55661.03883
1044981	29412	55661.03951
1049360	33791	55664.02213
1049361	33792	55664.02281
1053740	38171	55667.00544
1053741	38172	55667.00612
1058120	42551	55669.98875
1058121	42552	55669.98943
1062500	46931	55672.97205
1062501	46932	55672.97273
1066880	51311	55675.95536
1066881	51312	55675.95604
1073450	3401	55680.43032
1073451	3402	55680.43100
1077830	7781	55683.41362
1082210	12161	55686.39693
1086590	16541	55689.38024
1090970	20921	55692.36354

Continued on next page

Table 3 – continued from previous page

CIN	RCI	Date (MJD)
1095350	25301	55695.34685
1099730	29681	55698.33016
1104110	34061	55701.31346
1108490	38441	55704.29677
1116110	3281	55709.48690
1120490	7661	55712.47021
1120491	7662	55712.47089
1124870	12041	55715.45352
1124871	12042	55715.45420
1129250	16421	55718.43682
1129251	16422	55718.43750
1133630	20801	55721.42013
1133631	20802	55721.42081
1138010	25181	55724.40344
1138011	25182	55724.40412
1142390	29561	55727.38674
1142391	29562	55727.38742
1146770	33941	55730.37005
1146771	33942	55730.37073
1151150	38321	55733.35335
1155530	42701	55736.33666

5.4 Reaction Wheel Zero Crossings

The cadences occurring during reaction wheel zero crossings are listed in Table 4.

Table 4: Zero crossing events in Q9, defined as the time from first to last zero crossing in the event, rounded to the nearest cadence. The corresponding Short Cadence numbers can be found in the Data Release 12 Supplement. CIN = cadence interval number, RCI = relative cadence index.

Event no.	MJD start	MJD end	CIN start	CIN end	RCI start	RCI end
1	55643.122	55643.203	34340	34344	104	109
3	55649.109	55649.150	34633	34635	397	399
4	55658.059	55658.099	35071	35073	835	837
5	55669.971	55670.033	35654	35657	1418	1421
6	55675.958	55675.999	35947	35949	1711	1713
7	55683.417	55683.498	36312	36316	2076	2080

5.5 Downlink Earth Point

The cadences occurring during spacecraft Earth Point are listed in Tables 8 and 9. They are also flagged in the FITS files at MAST in the quality flag column (see the Kepler Archive Manual).

5.6 Manually Excluded Cadences

There were no manually excluded cadences in Q9. Any cadences that are manually excluded will be indicated in the quality flag column in the FITS files at MAST.

5.7 Incomplete Apertures Give Flux and Feature Discontinuities at Quarter Boundaries

No changes from the Data Characteristics Handbook.

5.8 Argabrightening

The cadences affected by Argabrightening events are listed in Tables 5 and 6 for LC and SC respectively. Argabrightenings listed in these tables were detected using the same algorithm as the pipeline, but with a threshold of 10 T_{MAD} instead of the pipeline threshold of 100. Only events exceeding the pipeline threshold are flagged in the FITS files

Table 5: Q9 LC Argabrightening Events with amplitude $T_{\mathrm{MAD}} > 10$, and occurring on a number of channels $T_{\mathrm{MCE}} > 10$. The columns are (1) Cadence interval number (CIN) for Argabrightening cadences, (2) Relative cadence index (RCI) for Argabrightening cadences, (3) Argabrightening cadence mid-times (MJD), (4) Mean Argabrightening statistic over channels included in the Argabrightening event $< S_{\mathrm{Arg}} >_{\mathrm{FPA}}$, (5) Number of channels exceeding threshold for this cadence ($\mathrm{N_{chan}}$), (6) Number of channels exceeding the default pipeline threshold for this cadence ($\mathrm{N_{pipe}}$).

CIN	RCI	Mid-Times (MJD)	$< S_{\mathrm{Arg}} >_{\mathrm{FPA}}$	$\mathrm{N_{chan}}$	$\mathrm{N_{pipe}}$
34287	51	55642.03864	7.3	20	0
34300	64	55642.30428	14.3	68	0
34499	263	55646.37056	16.5	65	0
34500	264	55646.39100	19.0	74	0
34546	310	55647.33094	37.5	79	0
34644	408	55649.33344	494.0	80	79
35095	859	55658.54899	47.4	79	0
35169	933	55660.06108	52.8	79	0
35222	986	55661.14406	494.4	80	79
35705	1469	55671.01349	24.7	79	0
35933	1697	55675.67235	12.5	56	0
36054	1818	55678.14482	24.1	79	0
36214	1978	55681.41420	19.2	73	0
36249	2013	55682.12937	10.0	48	0
36688	2452	55691.09972	8.0	21	0
36788	2552	55693.14309	21.8	78	0
36950	2714	55696.45333	11.0	50	0
37020	2784	55697.88368	99.2	80	42
37327	3091	55704.15680	19.8	76	0
37362	3126	55704.87197	488.8	80	79
37415	3179	55705.95496	16.9	73	0
37636	3400	55710.47078	2463.3	80	80
37783	3547	55713.47452	4251.2	80	80
37827	3591	55714.37360	9.1	31	0
37941	3705	55716.70303	6.8	18	0
37945	3709	55716.78477	7.4	22	0
37946	3710	55716.80520	6.6	16	0
38069	3833	55719.31853	14.3	57	0
38070	3834	55719.33897	57.0	80	1
38071	3835	55719.35940	30.5	76	0
38072	3836	55719.37983	11.9	45	0
38073	3837	55719.40027	10.4	35	0
38074	3838	55719.42070	22.1	74	0

Continued on next page

Table 5 – continued from previous page

CIN	RCI	Mid-Times (MJD)	$< S_{\mathrm{Arg}} >_{\mathrm{FPA}}$	$\mathrm{N_{chan}}$	$\mathrm{N_{pipe}}$
38085	3849	55719.64547	342.4	80	78
38138	3902	55720.72845	43.3	79	7
38170	3934	55721.38233	226.9	31	28
38375	4139	55725.57122	14.8	66	0
38654	4418	55731.27219	26.1	55	0
38694	4458	55732.08954	3.3	11	0
38721	4485	55732.64124	21.6	78	0
38736	4500	55732.94775	330.3	80	79

Table 6: Q9 SC Argabrightening Events with amplitude $T_{\mathrm{MAD}} > 10$, and occurring on a number of channels $T_{\mathrm{MCE}} > 10$. The columns have the same meanings as Table 5. Note consecutive detections of the largest events. A horizontal line separates the three months of the quarter; the relative cadence index (RCI) is reset at the start of each month.

CIN	RCI	Mid-Times (MJD)	$< S_{\mathrm{Arg}} >$	$\mathrm{N_{chan}}$	$\mathrm{N_{pipe}}$
1017089	1520	55642.04170	7.0	18	0
1017485	1916	55642.31143	13.7	60	0
1023431	7862	55646.36137	17.0	64	0
1023471	7902	55646.38861	18.9	71	0
1024857	9288	55647.33265	34.2	80	2
1027804	12235	55649.33991	446.8	80	80
1027805	12236	55649.34059	14.6	27	5
1041328	25759	55658.55138	44.3	80	14
1043543	27974	55660.06006	51.3	80	22
1045124	29555	55661.13691	386.9	80	80
1045125	29556	55661.13759	71.0	80	49
1059633	44064	55671.01928	24.4	80	0
1066470	50901	55675.67610	8.1	23	0
1070098	49	55678.14720	23.9	77	0
1074886	4837	55681.40841	20.3	77	0
1075938	5889	55682.12494	10.3	41	0
1089104	19055	55691.09257	9.0	23	0
1092117	22068	55693.14479	9.3	32	0
1092118	22069	55693.14547	9.0	29	0
1099078	29029	55697.88607	95.4	80	73
1099079	29030	55697.88675	11.5	39	0
1108296	38247	55704.16463	6.7	15	0
1108297	38248	55704.16531	7.8	18	0
1109339	39290	55704.87504	460.0	80	80
1109340	39291	55704.87572	57.8	80	28
1110911	40862	55705.94576	11.0	45	0
1117554	4725	55710.47044	275.7	80	80
1117555	4726	55710.47112	14.3	55	0
1117556	4727	55710.47180	793.2	80	74
1117557	4728	55710.47249	1059.3	74	70
1117558	4729	55710.47317	183.4	24	23
1121964	9135	55713.47418	212.4	80	80
1121965	9136	55713.47486	810.8	80	80
1121966	9137	55713.47554	991.3	80	80

Continued on next page

Table 6 – continued from previous page

CIN	RCI	Mid-Times (MJD)	$< S_{\mathrm{Arg}} >$	$\mathrm{N_{chan}}$	$\mathrm{N_{pipe}}$
1121967	9138	55713.47622	936.5	80	80
1121968	9139	55713.47691	565.1	80	80
1121969	9140	55713.47759	319.4	80	80
1121970	9141	55713.47827	80.3	80	49
1121971	9142	55713.47895	42.1	80	13
1121972	9143	55713.47963	6.4	16	0
1123291	10462	55714.37803	9.2	34	0
1126709	13880	55716.70610	7.4	22	0
1131031	18202	55719.64990	157.4	80	79
1131032	18203	55719.65058	184.7	70	61
1132612	19783	55720.72675	19.5	75	0
1132613	19784	55720.72743	14.8	14	9
1133572	20743	55721.38062	69.5	29	25
1133573	20744	55721.38131	89.9	25	22
1133574	20745	55721.38199	44.6	16	12
1139724	26895	55725.57088	14.4	65	0
1148103	35274	55731.27798	25.7	55	7
1150095	37266	55732.63477	19.8	73	0
1150541	37712	55732.93855	14.2	61	0
1150542	37713	55732.93923	123.3	80	78
1150543	37714	55732.93991	94.7	80	73
1150544	37715	55732.94060	48.8	80	20
1150545	37716	55732.94128	25.9	76	2
1150546	37717	55732.94196	5.9	11	0

5.9 Background Time Series

The background flux time series for Q9 is shown in Figure 6. We note that due to the presence of faint stars in the pixels used to measure the background flux, there is typically a small over-estimation in the background flux. For very faint targets ($K_p > 18$) this can result in occasional negative flux values in the time series. For brighter targets this has a negligible affect. If this is a concern, users are advised to add the background time series (provided in the FITS files) back to the flux time series, and perform their own background subtraction using appropriate pixels from the target pixel files (see the Kepler Archive Manual for more information).

5.10 Pixel Sensitivity Droputs

Pixel sensitivity dropouts, such as shown as a red curve in Figure 7, are thought to be caused by cosmic rays hits onto the detector. The PDC module now identifies and fixes most of these in the PDC light curves as shown as a blue curve in Figure 7. Cadences where these have been identified are flagged in the FITS files (see the Archive Manual for more details).

5.11 Short Cadence Requantization Gaps

No changes from the Data Characteristics Handbook.

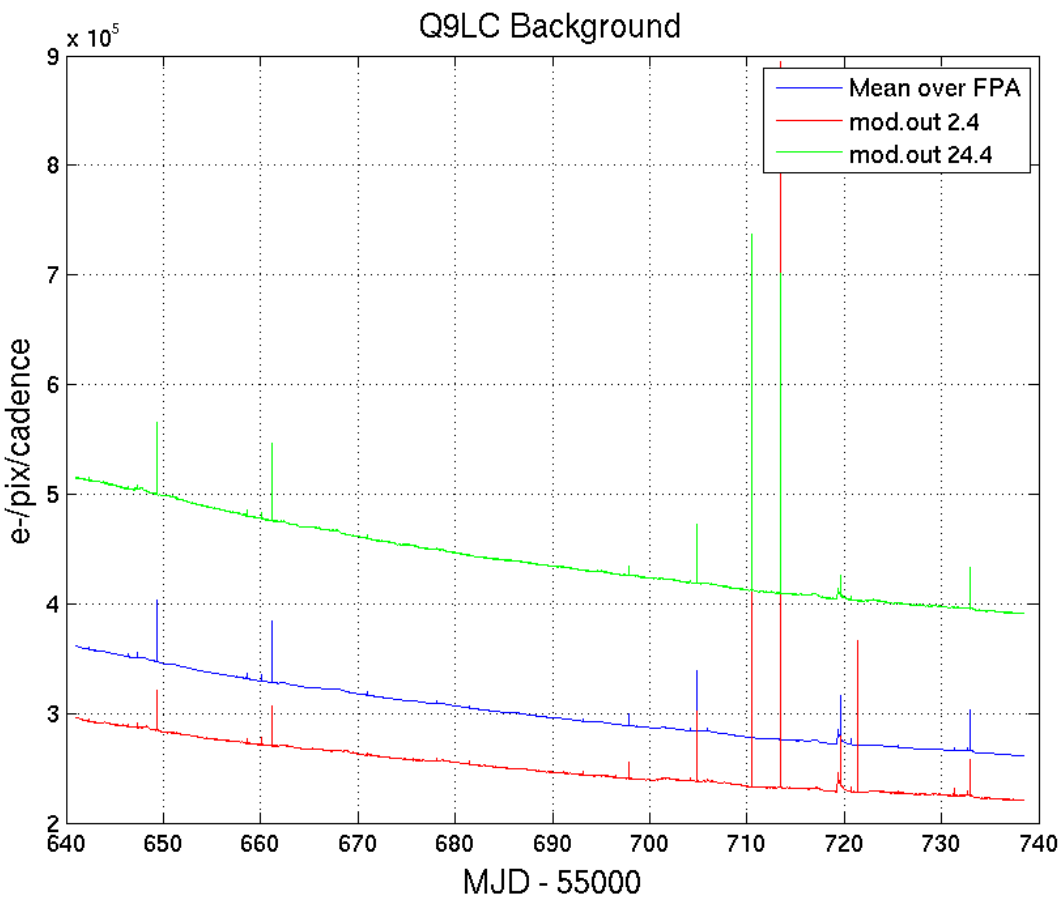

Figure 6: The background flux time series for Q9 showing the average over all the modules, and two individual modules. The focal plane average is shown in blue, modout 2.4 in red and modout 24.4 in green. The narrow spikes are Argabrightening events not identified as bad cadences by the Pipeline.

5.12 Spurious Frequencies in SC Data

A new list of spurious short cadence frequencies was supplied to us by Andrzej Baran of Cracow Pedagogical University. These were not vetted internally by the Kepler Science Office. These frequencies have been added to Table 7.

Table 7: Table of additional possible spurious frequencies detected in the SC data. Users are advised to check detections against this list and the list in the Kepler Data Characteristics Handbook, and report additional spurious frequencies. Labels: RW = reaction wheel passive thermal cycle associated with momentum cycle. RWTH = Reaction wheel housing temperature controller thermal cycling (believed not to be a problem from Q3 onward). U = unknown. LC are related to the long cadence read-out frequency. Narrow lines are defined as $\nu/\Delta\nu > 50$, broad lines as $\nu/\Delta\nu < 50$.

frequency	frequency	period	period	period	period			
μHz	d^{-1}	s	min	hr	d	Label	Width	Note
3.9	0.33	112320.00	1872.000	72.000	3.00	RW	?	
86.8	7.50	11520.00	192.000	3.2000	0.133	RWTH	broad	
196.5	16.98	5088.34	84.806	1.4134	0.05889		broad	[6]

Continued on next page

Table 7 – continued from previous page

frequency	frequency	period	period	period	period			
μHz	d^{-1}	s	min	hr	d	Label	Width	Note
242.5	20.95	4124.11	68.735	1.1456	0.04773		very broad	[5]
290.0	25.06	3448.28	57.471	0.9579	0.03991	U1	broad	
340.0	29.38	2941.18	49.020	0.8170	0.03404	U2	broad	
344.9	29.80	2899.33	48.322	0.8054	0.03356		very broad	[4]
360.0	31.10	2777.78	46.296	0.7716	0.03215	U3	narrow	
362.8	31.35	2755.98	45.933	0.7656	0.03190		very broad	[3]
365.8	31.61	2733.31	45.555	0.7593	0.03164		broad	[2]
370.4	32.00	2700.00	45.000	0.7500	0.03125	U4	narrow	
421.0	36.37	2375.30	39.588	0.6598	0.02749	split U5-U8	narrow	
566.4	48.94	1765.56	29.426	0.4904	0.02043	1/LC	narrow	
1132.8	97.87	882.78	14.713	0.2452	0.01022	2/LC	narrow	
1680.1	145.15	595.25	9.921	0.1653	0.00689			[10]
1699.2	146.81	588.52	9.809	0.1635	0.00681	3/LC	narrow	
1803.4	155.81	554.52	9.242	0.1540	0.00642			[1]
1825.2	157.7	547.88	9.131	0.1522	0.00634			[9]
2092.6	180.80	477.88	7.964	0.1327	0.00553			[8]
2265.6	195.74	441.39	7.357	0.1226	0.00511	4/LC	narrow	
2832.0	244.68	353.11	5.885	0.0981	0.00409	5/LC	narrow	
3398.3	293.62	294.26	4.904	0.0817	0.00341	6/LC	narrow	
3964.7	342.55	252.22	4.204	0.0701	0.00292	7/LC	narrow	
4428.4	382.61	225.82	3.763	0.0627	0.00261			
4451.4	384.6	224.65	3.744	0.0624	0.00260			[11]
4531.1	391.49	220.70	3.678	0.0613	0.00255	8/LC	narrow	
4995.8	431.64	200.17	3.336	0.0556	0.00232			[1]
5016.2	433.4	199.35	3.323	0.0554	0.00231			[11]
5097.5	440.43	196.17	3.270	0.0545	0.00227	9/LC	narrow	
5564.4	480.76	179.72	2.995	0.0499	0.00208		broad	[7]
5584.5	482.5	179.07	2.984	0.0497	0.00207			[11]
5663.9	489.36	176.56	2.943	0.0490	0.00204	10/LC	narrow	
6230.3	538.30	160.51	2.675	0.0446	0.00186	11/LC	narrow	
6796.7	587.23	147.13	2.452	0.0409	0.00170	12/LC	narrow	
7024.0	606.87	142.37	2.373	0.0395	0.00165	U5	narrow	[12,16]
7363.1	636.17	135.81	2.264	0.0377	0.00157	13/LC	narrow	
7444.0	643.16	134.34	2.239	0.0373	0.00155	U6	narrow	[13,16]
7865.0	679.54	127.15	2.119	0.0353	0.00147	U7	narrow	[14,16]
7929.5	685.11	126.11	2.102	0.0350	0.00146	14/LC	narrow	
8286.0	715.91	120.69	2.011	0.0335	0.00140	U8	narrow	[15,16]
8495.9	734.04	117.70	1.962	0.0327	0.00136	15/LC	narrow	

[1] Noted in Q7.
[2] 31.52–31.70 d^{-1}, and harmonic at 63.38 c/d.
[3] 31.17–31.53 d^{-1}, noted in Q7.
[4] 29.46–30.14 d^{-1}.
[5] Noted in Q7.
[6] Broad harmonics at 33.95 d^{-1} and 50.92 d^{-1}.
[7] 480.82 d^{-1} in Q6.3, 480.69 d^{-1} in Q7.1, 480.59 d^{-1} in Q7.2, 480.49 d^{-1} in Q7.3.
[8] Noted at 180.45 d^{-1} in Q6.3.
[9] Also noted at 157.97 d^{-1}.
[10] Also noted at 145.36 d^{-1}.
[11] 384.6 d^{-1}, 433.4 d^{-1} and 482.3 d^{-1} have the same separation as the 1/LC artifacts but are shifted from the LC comb by about 6.98 d^{-1}.

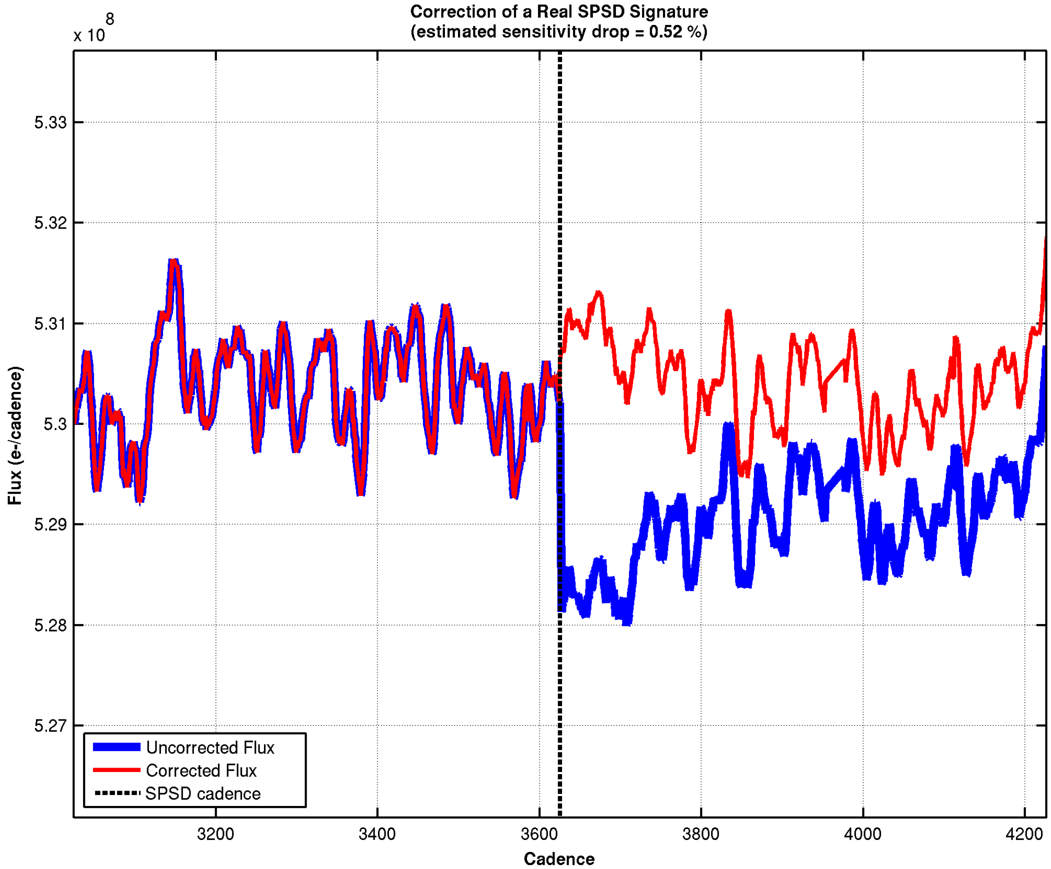

Figure 7: The effect of a sudden pixel sensitivity drop (SPSD). The black dashed line shows the cadence where the SPSD occurs. The blue curve follows the PA data, while the red line shows the PDC correction of the SPSD.

[12] Evolved to 609.29 d^{-1} by Q7.1.
[13] Evolved to 645.14 d^{-1} by Q7.3.
[14] Evolved to 680.73 d^{-1} by Q7.3.
[15] Evolved to 716.34 d^{-1} by Q7.3.
[16] These artifacts create a comb with separation 35.7 d^{-1}.

5.13 Propagation of Uncertaintities

No changes from the Data Characteristics Handbook.

5.14 LDE Out of Sync

During Q9 the Local Detector Electronics (LDE) became out of sync after the second Earth-point. This resulted in the first 30 short cadences not being processed. This is the first instance of this anomaly and is not flagged in the FITS files available from MAST. The cadences where this occurs are listed in Table 9. In the FITS files they have no data in the flux columns.

5.15 Anomaly Summary Tables

The full lists of affected cadence interval numbers (CINs) for Tables 8 and 9 below are included in the Data Release 12 Supplement. The Argabrightening Events listed here are those identified by the pipeline and are for completeness only. Those identified at a lower threshold are shown in Section 5.8 and should be considered the most complete list available

Table 8: Q9 LC Anomaly Summary Table

LC CIN start	LC CIN end	Anomaly Type	Note
36019	36052	EARTH POINT	
37448	37478	EARTH POINT	
38070	38070	ARGABRIGHTENING	
38138	38138	ARGABRIGHTENING	

Table 9: Q9 SC Anomaly Summary Table

SC CIN start	SC CIN end	Anomaly Type	Note
1018669	1018680	COARSE POINT	
1023079	1023090	COARSE POINT	
1027459	1027470	COARSE POINT	
1031839	1031850	COARSE POINT	
1036219	1036230	COARSE POINT	
1040599	1040610	COARSE POINT	
1044979	1044990	COARSE POINT	
1049359	1049370	COARSE POINT	
1053739	1053750	COARSE POINT	
1058119	1058130	COARSE POINT	
1062499	1062510	COARSE POINT	
1066879	1066890	COARSE POINT	
1073449	1073460	COARSE POINT	
1077829	1077840	COARSE POINT	
1082209	1082220	COARSE POINT	
1086589	1086600	COARSE POINT	
1090969	1090980	COARSE POINT	
1095349	1095360	COARSE POINT	
1099729	1099740	COARSE POINT	
1104109	1104120	COARSE POINT	
1105750	1105779	COARSE POINT	
1108489	1108500	COARSE POINT	
1112830	1112859	LDE OUT OF SYNC	See Section 5.14
1116109	1116120	COARSE POINT	
1120489	1120500	COARSE POINT	
1124869	1124880	COARSE POINT	
1129249	1129260	COARSE POINT	
1132612	1132613	ARGABRIGHTENING	
1133629	1133640	COARSE POINT	
1138009	1138020	COARSE POINT	
1142389	1142400	COARSE POINT	
1146769	1146780	COARSE POINT	
1151149	1151160	COARSE POINT	

Continued on next page

Table 9 – continued from previous page

SC CIN start	SC CIN end	Anomaly Type	Note
1155529	1155540	COARSE POINT	

6 Time and Time Stamps

No changes from the Data Characteristics Handbook.

7 Ensemble Cotrending Basis Vectors

NEW: CBVs from clean reference ensembles are now generated directly by Pipeline instead of by the offline tool used for Q1-Q8; users are provided with a quantitative measure of goodness of CBVs.

7.1 Introduction

We infer the presence of systematic errors in the Kepler flux time series from the correlations observed between them, since we do not expect the stars themselves to have correlated signals. These correlations can be represented as linear combinations of orthonormal functions, called cotrending basis vectors (CBVs), which in some sense represent most of the correlated features in a reference ensemble of flux time series for a given Quarter and output channel (mod.out). Light curve correction can be performed by finding the 'best' fit of these CBVs to a light curve, and removing the fit. The CBVs of a reference ensemble of highly correlated stars are provided so that users can perform their own systematic error removal without having particular knowledge of proprietary targets, using their own tools or those supplied by the GO office.

In earlier releases, PDC used engineering telemetry and local image motion polynomials derived by the Pipeline itself to remove systematic errors from target flux time series. We have also generated CBVs for these earlier Releases and delivered them to MAST, even though the Pipeline version in use at the time did not use CBVs.

User systematic error correction using the CBVs may be preferable to PDC data, or necessary, since:

1. An ensemble of flux time series is potentially a more complete representation of the trends than engineering telemetry and image motion,
2. Users can decide how many CBVs to use to fit systematic trends,
3. Users can use the CBVs with non-least squares fitting methods such as MAP (Jenkins et al., 2011) or the lasso (Tibshirani 1994; `http://www-stat.stanford.edu/~tibs/ftp/lasso.ps`).
4. Users of target pixel files who generate their own flux time series will need to do their own systematic error removal, though it is up to the user to understand whether a set of CBVs which well represent systematic errors in uncorrected Pipeline flux time series is also a good representation of their systematic errors given differences in aperture size and their method of extracting a flux time series from the pixels.

7.2 Generation of CBVs

The method for CBV generation in Pipeline 8.0 will be detailed in a future release of the Kepler Data Analysis Handbook. Briefly, the method:

1. Removes the median from each uncorrected flux time series,
2. Normalizes each median-removed flux time series by its RMS,
3. Calculates the correlation between RMS-normalized flux time series and selects the 50% most correlated stars as the reference ensemble,
4. Estimates the intrinsic variability of each star in the reference ensemble by removing a 3rd order polynomial fit, and calculating the standard deviation of the detrended light curve normalized by flux uncertainty (roughly, read and shot noise). Stars with a normalized detrended standard deviation 30% greater than the median value on a particular mod.out are rejected from the reference ensemble.
5. Performs a Singular Value Decomposition (SVD) of the median-removed, median-normalized, variable-cleaned reference ensemble. The CBVs in this case are the SVD principal components, though a more general nomenclature is used to allow non-SVD approaches to be used in the future. Users are provided with the leading 16 components. Users are also provided with the SVD principal values for each mod.out, to inform their decision about how many components to use in their fits.

A good CBV will be representative of many light curves instead of just one or a few. To assess this, we use the robust least-squares fit coefficients for the stars in the reference ensemble, which are generated by the Pipeline as part of the Bayesian Maximum A Posteriori (MAP) method. It can be shown that, for an orthonormal basis, the distribution of fit coefficients is equivalent to the distribution of coefficients that describe the linear combination of light curves which make up each CBV. So a CBV for which a few stars have unusually large fit coefficients is also a CBV which is mostly made up of a linear combination of those few stars' uncorrected light curves, and should not be used. The quantitative metric used to determine the 'quality' of a given CBV is the relative entropy h_i of the probability distribution function p_i of the robust fit coefficients θ_i of the ith CBV, relative to a Gaussian of the same standard deviation:

$$h_i = -\int p_i(\theta_i) \ln p_i(\theta_i) d\theta_i - H_0(\sigma) \tag{1}$$

where

$$H_0(\sigma) = \frac{1 + \ln 2\pi}{2} + \ln \sigma \tag{2}$$

is the entropy of a Gaussian of standard deviation σ.

7.3 Using CBVs

To use the CBVs for least-squares fitting, subtract the median uncorrected flux from the uncorrected flux time series of interest and divide by the median. Since the basis is orthonormal, the linear least-squares fit coefficient of the nth CBV is simply the inner product of the median-removed, median-normalized uncorrected flux time series with the nth CBV. Subtract the fit to get the corrected (median-removed, median-normalized) flux time series.

Convenient CBV tools, including robust fitting and time window exclusion, are provided by the Guest Observer Office as part of PyKE and are available from `http://keplergo.arc.nasa.gov/ContributedSoftwarePyKEP.shtml`. The site provides instructions on how to install the software and specific instructions on how to use the tools to fit the CBVs. Users should note that, unlike the SOC Pipeline, these tools do not include scalar amplitude corrections for the fraction of target flux captured in the optimal aperture, or the fraction of the total flux in that aperture which is from the target star and not from its neighbors or unresolved background objects. These quantities (the flux fraction in aperture and the crowding metric, respectively) are available for the optimal apertures computed by the SOC as keywords in the FITS files at MAST, or for quarters not yet processed by the SOC 8.0 pipline, from the data search page at MAST.

Users will need to use **at least** the first two CBVs, since the generation method mixes a constant offset with the strongest non-constant component instead of strictly enforcing a constant first or second component; using only the first component would be like attempting a linear fit with a constant or slope term, but not both. Figure 8 shows that 8 or fewer components generally capture most of the systematic error for all mod.outs, though 16 are provided if users wish to make their own decisions.

7.4 Quality of CBVs

Generally, CBVs with a relative entropy < -1 should not be used. In Release 12, *no* strong CBVs (indices 1–4) are bad, and only 13/320 weak CBVs (indices 5–8) are bad, as shown in Table 10

7.5 Cautions

1. Channels with strong Moiré pattern drift (see the Instrument Handbook, Section 6.7) have less stable weak (>5th order) CBVs, in the sense that random constant perturbations of the reference ensemble give results which differ by significantly more than a constant offset. For good channels, the stability of components to the 10th order has been verified.

2. While most variable stars have been removed from the reference ensemble, it is still possible that some weaker CBVs have been influenced by noisy stars or by variable stars whose signal slips through the

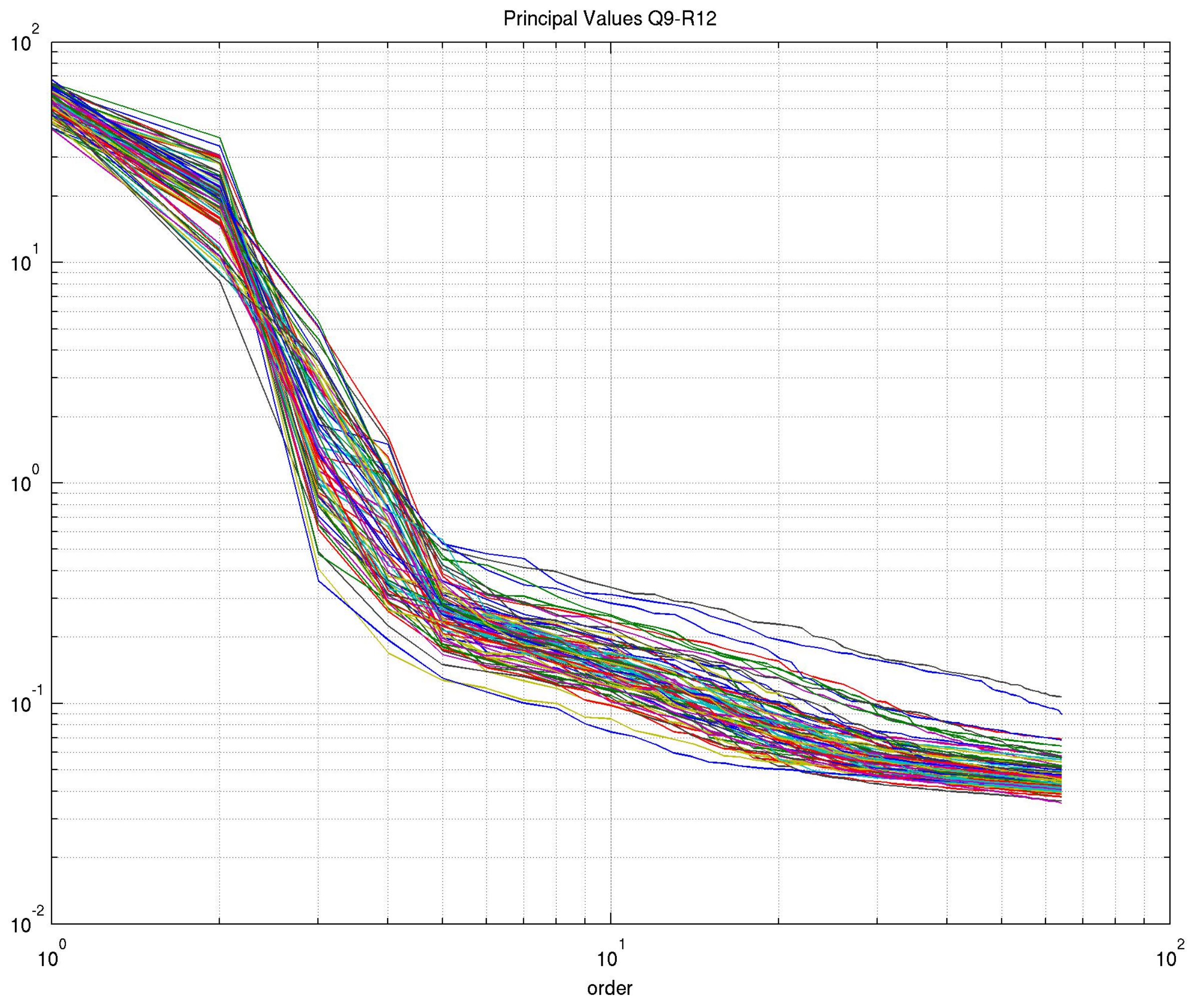

Figure 8: Principal values of SVD-extracted cotrending basis vectors for Quarter 9, Release 12, showing that most of the systematic error can be accounted for by the first 8 or fewer components. Users are provided with the first 16 components for each mod.out, as well as the principal values which are plotted in this Figure.

detrending filter. Users should exclude bad CBVs from their fits, and in general be cautious about reporting results with the same period and phase as one of the basis vectors used in the fit.

3. No vetting of CBVs 9-16 has been done, as the Pipeline does not generate the robust fit coefficients for these CBVs since they are not used in MAP. These CBVs should be used with extreme caution.

7.6 Obtaining CBVs

The cotrending basis vectors can be downloaded in FITS files from the MAST website (`http://archive.stsci.edu/kepler`) separate from your data download. There is one file per quarter containing 84 extensions, one for each channel. Each extension contains 16 basis vectors along with the cadence and MJD of the observations. The cadences found in the basis vector file match the number of cadences in the light curve file for that quarter. To ensure that your data was processed through the same version of the pipeline as the stars used to create the basis vectors, the keyword DATA_REL in your data's light curve file should match that found in the basis vector file. A new basis vector file will be provided each

Table 10: Bad cotrending basis vectors, with relative entropy < -1.

channel	module	output	CBV #	Relative Entropy
11	4	3	5	-1.2889
23	8	3	6	-2.0120
36	11	4	7	-1.5329
36	11	4	8	-1.0829
47	14	3	5	-1.0295
58	17	2	5	-1.4416
63	18	3	5	-1.2964
63	18	3	8	-1.1311
67	19	3	7	-1.3937
71	20	3	6	-1.2322
72	20	4	7	-1.0951
83	24	3	6	-1.1271
84	24	4	6	-1.0749

time the data are reprocessed.

8 Contents of Supplement

The Supplement is available as a full package (DataReleaseNotes12SupplementFull.tar), which contains the files described below.

Pipeline Instance Detail Reports

```
q9-lc-tad+cal+pa+tad+pa-r8.0-ksop-900-as-run-pipeline-instance-detail-report-111013.txt
q9-lc-pdc-through-ppa-r8.0-ksop-900-as-run-pipeline-instance-detail-report-111017.txt
q9-sc-m1-m2-m3-r8.0-ksop-900-as-run-pipeline-instance-detail-report-111027.txt
q9-scm3-pa-pdc-r8.0-ksop-900-as-run-pipeline-instance-detail-report-111027.txt
```

Data Anomaly Types

```
DataAnomalyTypes_Q9LC_LC_PID5542.txt
DataAnomalyTypes_Q9SCM1_SC_PID5562.txt
DataAnomalyTypes_Q9SCM2_SC_PID5562.txt
DataAnomalyTypes_Q9SCM3_SC_PID5587.txt
```

Mod.out central motion

```
Q9LC_central_column_motion.txt
Q9LC_central_row_motion.txt
```

Thermal Telemetry

```
Q9_LDE_averageBoardTemp.txt
Q9_TH12LVAT_MJD_gap.txt
Q9_TH1RW34T_MJD_gap.txt
```

Background Flux Time Series

```
Q9LC_background.txt
Q9SCM1_background.txt
Q9SCM2_background.txt
Q9SCM3_background.txt
```

Argabrightening Detections

```
Q9LC_LC_ArgAgg_Summary.txt
Q9SCM1_SC_ArgAgg_Summary.txt
Q9SCM2_SC_ArgAgg_Summary.txt
Q9SCM3_SC_ArgAgg_Summary.txt
```

Out of Fine Point Cadence Lists

```
Q9LC_LC_isNotFinePoint.txt
Q9SCM1_SC_isNotFinePoint.txt
Q9SCM2_SC_isNotFinePoint.txt
Q9SCM3_SC_isNotFinePoint.txt
```

Zero Crossing Events

```
Q9LC_ZeroCrossings.txt
```

```
Q9SCM1_ZeroCrossings.txt
Q9SCM2_ZeroCrossings.txt
Q9SCM3_ZeroCrossings.txt
```

Cotrending Basis Vector Entropy and Principal Values

```
Q9_R12-principal_values_L-0.50_SVDO-8.txt
Q9-R12_DAWG_CBV_entropy_bad_CBVs.txt
```

8.1 Short Supplement Package

The Supplement also contains a short package suitable for emailing (DataReleaseNotes12SupplementShort.tar). The small package does not contain the following files:

```
Q9LC_background.txt
Q9SCM1_background.txt
Q9SCM2_background.txt
Q9SCM3_background.txt

Q9LC_central_column_motion.txt
Q9LC_central_row_motion.txt

Q9_TH12LVAT_MJD_gap.txt
Q9_TH1RW34T_MJD_gap.txt
```

9 References

In addition to those found in the Data Characteristics Handbook:

Jenkins, J. M., Smith, J. C., Tenenbaum, P., Twicken, J. D., and Van Cleve, J., 2011, in Advances in Machine Learning and Data Mining for Astronomy (eds. Way, M., Scargle, J., Ali, K., Srivastava, A.,), Chapman and Hall/CRC Press

Shupe, D. L., et al., 2005, Astronomical Data Analysis Software and Systems XIV, 347, 491

Skrutskie, M. F., et al., 2006, The Astronomical Journal, 131, 1163

A SOC-8.0 Improvements

The files which make up the Q9 data release are the first to have been made using the the SOC 8.0 pipeline. There are a number of changes and improvements related to this release which are summarised here.

A.1 Improvements to PDC

The Presearch Data Conditioning (PDC) module for long cadence data has been completely redesigned for the Q9 data release. The old PDC would occasionally distort astrophysical signals. In the new PDC this problem has been significantly reduced.

For details on the new PDC data, see the Archive Manual. Also, please note that even though new PDC keywords are populated in the FITS headers of short cadence data, the improvements to PDC do not apply to short cadence light curves.

A.2 New FITS header keywords

Many new keywords can be found in the headers of the version 2.1 FITS files. For details of these, please see the Archive Manual. 3, 6 and 12-h CDPP, flux fraction, and crowding have been extracted from the FITS files and are available via the data search page at MAST. Values are also provided for all past quarters.

A.3 Improved World Coordinates in Full Frames Images

The Q9 Full Frame Images (FFIs) have an improved world coordinate system (WCS) solution when compared to older FFIs. The new FFIs use a non-linear WCS based on the simple imaging polynomal (SIP) convensions of Shupe et al. (2005). The transformations from sky coordinates to pixel coordinates is calculated by the Kepler pipeline for each long cadence but not for the FFIs. As a work-around we use the nearest good long cadence to calculate the coordinate transformations for the FFI (typically the last long cadence of the month). The WCS solution is typically accurate to 0.1 pixels (0.4 arcseconds) when compared to the 2MASS point source catalogue (Skrutskie et al. 2006). For full details, see the Archive Manual.

There is also a minor change in our release policy. Starting with Q9, the FFIs will be released quarterly with the data release instead of monthly as before because their calibration is now tied to the quarterly processing.

B A problem with the Target Pixel File World Coordinates

An issue has been identified with the FITS keywords which describe the World Coordinates in the target pixel files. The 1CDLT and 2CDLT keywords describe the right ascention and declination vectors. The values for these are switched in the target pixel cadence images but not in the aperture definition. This error is present in all Q0–9 target pixel files. It will be fixed for Q10 data.